Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ON MATRICES WHOSE ELEMENTS ARE INTEGERS.

By Oswald Veblen and Philip Franklin.

Introduction.

1. The purpose of this article is strictly expository. The aim is to set forth some of the theorems on matrices whose elements are integers. These theorems have applications in Analysis Situs* and the systematic treatment of them directly in terms of integers here given will no doubt be useful to students of that subject. While the closely allied algebraic theory is to be found in Bôcher's Introduction to Higher Algebra, and the matter here given is to some extent discussed in Muth's Elementartheiler and in Scott and Mathews' Determinants, there is no readily accessible treatment of the subject from the point of view here adopted.

2. The object of our study will be a matrix of α rows and β columns:

$$(1) \qquad E = \| \epsilon_i^{\ j} \| = \left\| \begin{matrix} \epsilon_1^{\ 1} & \epsilon_1^{\ 2} & \cdots & \epsilon_1^{\ \beta} \\ \epsilon_2^{\ 1} & \epsilon_2^{\ 2} & \cdots & \epsilon_2^{\ \beta} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \epsilon_\alpha^{\ 1} & \epsilon_\alpha^{\ 2} & \cdots & \epsilon_\alpha^{\ \beta} \end{matrix} \right\| .$$

The elements of E are integers. The term "integer" here includes negative integers and zero; but we shall assume that at least one element is different from zero.

Our definition of the *product* of two matrices $\| \epsilon_i^{\ j} \|$ and $\| \eta_i^{\ j} \|$ is:

$$(2) \qquad \| p_i^{\ j} \| = \| \epsilon_i^{\ j} \| \cdot \| \eta_i^{\ j} \| ,$$

where

$$(3) \qquad p_i^{\ j} = \sum_{k=1}^{k=\beta} \epsilon_i^{\ k} \cdot \eta_k^{\ j}.$$

The number of rows of the second matrix must be equal to the number of columns of the first; and the product has as many rows as the first matrix and as many columns as the second. If the matrices are square, the product will be square, and the determinant of the product will be equal to the product of the determinants of the factors.

The *inverse* of a square matrix A, of determinant unity, will be the matrix A^{-1} such that:

$$A^{-1} \cdot A = A \cdot A^{-1} = I,$$

---

* Cf. O. Veblen, Cambridge Colloquium Lectures on Analysis Situs.

where I denotes the identity matrix $\| \delta_i^j \|$, a square matrix with all the elements in the main diagonal $+1$ and all the remaining elements zeros. The element $\check{a}_i^j$ of A^{-1} will evidently be the cofactor of a_j^i in the determinant of A. A is restricted to be of determinant unity to insure the elements of the inverse matrix being integers.

Elementary Transformations.

3. Let us consider two types of transformations of E:

(*a*) *To replace each element of the rth row* (ϵ_r^j) *by the element* $(\epsilon_r^j + q\epsilon_s^j)$ *where* q *is either* $+1$ *or* -1 *and* $s \neq r$. This operation is described as adding the sth row to the rth row or subtracting the sth row from the rth row.

(*b*) *To add a column to or subtract it from another column.*

The operation (*a*) is equivalent to multiplying E on the left by a square matrix of α rows $A_0 = \| a_i^j \|$ in which all the elements are zeros except those of the main diagonal which are $+1$, and a_r^s which is q. For the expressions given by (3) for the elements of the product, i.e.,

$$p_i^j = \Sigma a_i^k \cdot \epsilon_k^j \tag{4}$$

reduce to the single term ϵ_i^j except when $i = r$; in which case they give the two terms:

$$\epsilon_r^j + q\epsilon_s^j.$$

That is, the operation (*a*) transforms E into $A_0 \cdot E$.

In like manner, the operation (*b*) corresponds to multiplying E on the right by a square matrix of β rows $B_0 = \| b_i^j \|$ in which all the elements are zeros except those of the main diagonal which are 1 and b_s^r which is q.

If the operation (*a*) be repeated n times, where n is a positive integer, the effect is an operation identical with (*a*) except that q is replaced by the integer n or $-n$. Correspondingly, the effect of multiplying the matrix A_0 by itself repeatedly is to change the element a_r^s to $\pm n$.

The inverse of A_0 if $a_r^s = \pm 1$ is the same matrix except that the sign of a_r^s is changed. Hence the inverse of an operation of type (*a*) is an operation of the same type. The determinant of A_0 is $+1$.

Similar statements hold with regard to the operation (*b*) and the matrix B_0.

4. *The operation of interchanging two rows of a matrix and changing the signs of all the elements of one of them can be expressed as a sequence of operations of type* (*a*). For if we add the rth row to the sth, then subtract the sth row of the resulting matrix from the rth, and finally add the rth row to the sth, the elements of the rth and sth rows (and the qth column) will be, successively:

$$(\epsilon_r^q, \epsilon_s^q); \qquad (\epsilon_r^q, \epsilon_r^q + \epsilon_s^q); \qquad (-\epsilon_s^q, \epsilon_r^q + \epsilon_s^q); \qquad (-\epsilon_s^q, \epsilon_r^q);$$

and the resulting matrix will thus be that obtained by changing the signs of the elements of the sth row and then interchanging the rth and sth rows.

In like manner, *the operation of interchanging two columns and changing the signs of the elements of one of them is expressible as a sequence of operations of type* (*b*).

5. In place of our two fundamental operations (*a*) and (*b*) we might have restricted ourselves to the operations:

(a') *To add a row to, or subtract it from, an adjacent row.*

(b') *To add a column to, or subtract it from, an adjacent column.*

We reduce the operation (*a*) to a sequence of operations (a') in the following manner. For definiteness, let us speak of the sth row as following the rth; in the reverse case "following" is to be replaced by "preceding" in the argument. Add each row from the rth to the $(s-1)$th inclusive to the next following row, beginning with the $(s-1)$th. Then, in the resulting matrix subtract each row from the $(r+1)$th to the $(s-1)$th inclusive from the following row, beginning at the $(r+1)$th. Next add each row from the $(r+1)$th to the $(s-2)$th inclusive to the following, beginning at the $(s-2)$th. Finally subtract each row from the rth to the $(s-2)$th from the following row, beginning at the rth.

If the reader will write out the expressions for the elements of the matrix in a single column and the rows affected (from the rth to the sth), he will find that the resulting matrix only differs from our original matrix in having its rth row added to its sth (or subtracted from it). To perform the inverse operation we need only repeat the process, subtracting the rth row from the $(r+1)$th in the second step, and adding it to the $(r+1)$th in the last step; the other operations remaining as before.

As a similar argument holds for steps (*b*) and (b'), if we replace rows by columns, we conclude that *the transformations built up from steps* (*a*) *and* (*b*) *are no more general than those built up from steps* (a') *and* (b').

Determinant Factors.

6. Consider the set of γ-rowed $(0 < \gamma \leqq \alpha,\ \gamma \leqq \beta)$ determinants which can be formed from E by omitting $\alpha - \gamma$ of the rows and $\beta - \gamma$ of the columns in all possible ways. The highest common factor (H. C. F.) of such a set of determinants, if the determinants are not all zero, is denoted by D_γ and is called the γth *determinant factor*[*] of E.

The determinant factors are unchanged when the matrix is operated on by transformations of type (*a*) *or* (*b*). For, consider the effect of an operation of type (*a*) which consists in adding the rth row to (or subtracting it from) the sth, on the γ-rowed determinants in question. All such determinants which do not contain elements from the sth row are obviously unaffected, while those that contain elements from both the rth row and the sth are not affected because of an elementary theorem on determinants. The remaining γ-rowed determinants, as A_γ, which contain elements from the sth row and not from the rth, are converted into determinants of the form $A_\gamma \pm A_\gamma'$ where A_γ' is the γ-rowed determinant obtained from A_γ by replacing the elements from the sth row by elements from the same columns of the matrix and from the rth row.

---

[*] Cf. Scott and Mathews' Determinants, p. 76.

The proof for operations of type (b) is similar.

7. The following theorem has an application in Analysis Situs: *If a matrix E is such that each column either consists entirely of zeros or contains just two elements different from* 0, *one* $+1$ *and the other* -1, *all the determinant factors of the matrix are* $+1$ *or* -1.

The theorem follows immediately from the definition of a determinant factor, if we observe that any γ-rowed determinant formed by striking out $(\alpha - \gamma)$ rows and $(\beta - \gamma)$ columns of the given matrix has either two, none or one element in each column different from zero. If no column is of the third type the determinant is zero, since the sum of all the elements in each column is zero. If there is a column of the third type we evaluate the determinant with reference to such a column and then evaluate the minor with reference to a column with a single non-zero element in the minor, and so on. In this way we either arrive finally at ± 1 for the value of the determinant, or else come to a minor with two or no non-zero elements in each column, in which case the determinant is zero.

Reduction to Normal Form.

8. Let us now consider a series of reductions of the matrix E which can be effected by transformations of types (a) and (b). If the first column consists entirely of zeros, add one of the other columns to it. Thus by a transformation of type (b) E is converted into a matrix E_1 which has at least one non-zero element in the first column. If the first element of the first column is zero, add a row which contains a non-zero element in the first column to the first row. Thus by a transformation of type (a) E_1 is converted into a matrix E_2 for which the element of the first row and column is not zero.

We shall now prove that if this non-zero element ϵ_1^1 is not a factor of all the elements of the matrix, we can, by a series of transformations of types (a) and (b), replace it by a numerically smaller element different from zero.

First, if one of the elements in the first column, ϵ^1_r, is not divisible by ϵ_1^1, upon adding the first row to (or subtracting it from) the rth a number of times equal to the largest integer in the quotient of ϵ^1_r by ϵ_1^1, an element numerically smaller than ϵ_1^1 is obtained in the first column and rth row. Then, on subtracting the rth row from (or adding it to) the first row the matrix is converted into one with a smaller non-zero element in place of ϵ_1^1. This has been done by a succession of operations of type (*a*). Similarly, if there were an element in the first row which did not contain ϵ_1^1 as a factor, transformations of type (*b*), strictly analogous to those of type (*a*) just described, could be set up which would reduce the numerical value of ϵ_1^1.

Second, if ϵ_1^1 is a factor of all the elements of the first row and first column, but is not a factor of the element in the rth row and sth column, ϵ_r^s, we proceed as follows. Upon subtracting the first column from (or adding it to) the sth $\epsilon_1^s/\epsilon_1^1$ times (transformations of type (*b*)), the first element in the sth column becomes zero, while the rth is still not divisible by ϵ_1^1, since it has been changed by a multiple of ϵ_1^1. If we now add the sth column to the first (an operation of type (*b*)), the element in the first row and column remains ϵ_1^1, while the rth element in the first column is now not divisible by ϵ_1^1. Hence we may replace ϵ_1^1 by a numerically smaller element by the method of the preceding paragraph.

If the element which replaces ϵ_1^1 is not a factor of all the elements of the matrix, it may be still further reduced by a repetition of the process described in the two paragraphs above. If this process be continued, we must arrive after a finite number of steps—the number being less than the absolute value of ϵ_1^1—at a matrix whose first element d_1 is a factor of all the others. When this point is reached, we may reduce all the elements in the first column except the first to zeros by operations of type (*a*), for we have merely to add the first row to (or subtract it from) any other row the number of times the first element of this row contains d_1. The elements of the first row, with the exception of the first, may be reduced to zeros by similar operations of type (*b*). It is evident that all the elements of the matrix thus obtained contain the first element as a factor.

Thus we arrive at a matrix E_3 in which the first element d_1 of the first column is the H. C. F. of all the elements of E_3 and in which all the other elements of the first row and of the first column are zero. By § 6, d_1 is the H. C. F. of all the elements of E.

9. Let $\overline{E}_3$ be the matrix obtained from E_3 by deleting its first row and first column. By § 8, $\overline{E}_3$ may be reduced to a matrix with a leading element which is the H. C. F. of all its elements, and having all the other elements of the first row and column zero.

As the transformations of types (*a*) and (*b*) which effect this reduction on $\overline{E}_3$ determine transformations of E_3 of the same type, which leave its first row and first column unchanged, we may reduce the matrix E_3 to a matrix E_4 in which the first element of the main diagonal, d_1, is the H. C. F. of all the elements of the matrix, the second element of the main diagonal, d_2, is the H. C. F. of all the elements except d_1, and all the remaining elements of the first two rows and first two columns are zero.

By a continuation of this process we arrive by a finite sequence of operations of types (a) and (b) at a matrix:

$$(5) \qquad E^* = \begin{vmatrix} d_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & d_r & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{vmatrix},$$

in which all the elements are zero except a sequence of elements d_i, $(0 < i \leqq r)$ common to the ith row and column and such that d_i is the H. C. F. of all d_j's such that $i \leqq j \leqq r$.

The d_i's may be positive or negative integers. We can make all except the last positive by a sequence of operations of type (a). For if d_i is negative, and we interchange the ith and rth rows, changing the sign of the elements in the rth, a permissible operation by § 4, and repeat the process, we arrive at a form in which d_i is positive and d_r has changed sign. We may thus obtain a form in which d_i $(i < r)$ is positive, and d_r will be positive or negative according as the number of negative signs in the form we started with was even or odd. We shall take this matrix, with at most one negative element, as the normal form E^* in the discussion which follows.

Each operation of type (a) amounts, according to § 3, to multiplying the matrix to which it is applied on the left by a square matrix of type A_0 of α rows, and each operation of type (b) amounts to multiplying the matrix to which it is applied on the right by a square matrix of type B_0 of β rows. Hence

$$(6) \qquad E^* = A \cdot E \cdot B,$$

where A is a product of matrices of type A_0 and B a product of matrices of type B_0. It is to be noted that the determinants of A and B are each $+1$.

Let us introduce the notation $D_1 = d_1$, $D_2 = d_1 \cdot d_2$, $\cdots$ $D_r = d_1 \cdot d_2 \cdot \cdots d_r$, and observe that D_γ $(0 < \gamma \leqq r)$ is the H. C. F. of all the γ-rowed determinants which can be formed by striking out $\alpha - \gamma$ rows and $\beta - \gamma$ columns from E^*. That is, referring to § 6, they are the successive determinant factors of E^*. Since E^* was derived from E by operations of types (*a*) and (*b*), they are also the determinant factors of E.

Since the D_i's are invariant under transformations of types (*a*) and (*b*), the d_i's, which are the quotients of successive D_i's, $(d_{i+1} = D_{i+1}/D_i)$, are also invariant under these transformations. They are called *invariant factors* or elementary divisors.*

The number r is also invariant under transformations of types (*a*) and (*b*) and is called the *rank* of the matrix E.

The Matrices of Transformation.

10. In the special case where E is a square matrix of α rows whose determinant is $+1$, equation (6) implies that the determinant of E^* is $+1$. Hence $r = \alpha$, and the numbers d_i must be $+1$.

We therefore have:

$$A \cdot E \cdot B = I, \tag{7}$$

in which A is a product of matrices of type A_0, B a product of matrices of type B_0 and I is the identity matrix. We may write (7) in the form

$$E = A^{-1} \cdot I \cdot B^{-1} = A^{-1} \cdot B^{-1}. \tag{8}$$

It is evident from § 2 that when $\alpha = \beta$ every matrix of type B_0 can be regarded also as one of type A_0, and the same is true of matrices inverse to those of types A_0 or B_0. As the above equation shows that E is equal to a product of such matrices, we have the theorem: *Any square matrix of determinant unity is expressible as a product of matrices which may be considered to be of type A_0, or to be of type B_0.*

Hence to multiply a matrix E of α rows and β columns on the left by a square matrix of α rows and determinant unity is equivalent to operating on E by a sequence of operations of type (*a*); *and to multiply E on the right by a square matrix of β columns and determinant unity is equivalent to operating on E by a sequence of operations of type* (*b*).

Also, since we may write (8) in either of the forms:

$$B \cdot A \cdot E = I \qquad \text{or} \qquad E \cdot B \cdot A = I, \tag{9}$$

it follows that *if E is a square matrix of determinant unity, it may be reduced to the form I by operations on rows only, or by operations on columns only.*

11. In the case of a general matrix E, we have from (6)

$$A \cdot E = E^* \cdot B^{-1}. \tag{10}$$

Since the determinant of B^{-1} is 1, the H. C. F. of the elements of its first row is 1. Hence the H. C. F. of the elements of the first row of the matrix $E^* \cdot B^{-1}$ is d_1. As a similar statement applies to the remaining rows, we have the theorem:

---

* We shall use the term invariant factor, following Bôcher, Introduction to Higher Algebra, pp. 269–70, since the term elementary divisor is sometimes used in another sense.

The matrix A has the property that the H. C. F. of the elements of the rth row of the matrix $A \cdot E$ is d_r, the rth invariant factor of E.

This suggests a method of building up A by means of the theorems:

(1) That for any set of integers ϵ_j^1 $(0 < j \leqq \alpha)$, a set of integers a_1^j $(0 < j \leqq \alpha)$ can be found which are relatively prime and such that

$$\sum_{j=1}^{j=\alpha} a_1^j \cdot \epsilon_j^1 = t_1$$

where t_1 is the H. C. F. of the α ϵ_j^1's; and

(2) That there exists a matrix A of determinant unity with the numbers a_1^j as the elements of its first row.

The derivation of equation (6) by this method is longer than that given in §§ 8 and 9 and is therefore omitted.

Diophantine Equations of the First Degree.

12. Consider the problem of finding the integral solutions of the following set of equations:

$$(11) \qquad \begin{array}{l} \epsilon_1^1 x_1 + \epsilon_1^2 x_2 + \cdots + \epsilon_1^\beta x_\beta = p_1, \\ \epsilon_2^1 x_1 + \epsilon_2^2 x_2 + \cdots + \epsilon_2^\beta x_\beta = p_2, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \epsilon_\alpha^1 x_1 + \epsilon_\alpha^2 x_2 + \cdots + \epsilon_\alpha^\beta x_\beta = p_\alpha. \end{array}$$

If X denotes the matrix of one column, and β rows

$$\left\| \begin{array}{c} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_\beta \end{array} \right\|,$$

and P a similar matrix with $p_1, p_2 \cdots p_\alpha$ as the elements of its one column and α rows, equations (11) may be written

$$(12) \qquad E \cdot X = P.$$

But from (6) we have

$$(13) \qquad E = A^{-1} \cdot E^* \cdot B^{-1},$$

and consequently

$$(14) \qquad A^{-1} \cdot E^* \cdot B^{-1} \cdot X = P,$$

or

$$(15) \qquad E^* \cdot B^{-1} \cdot X = A \cdot P.$$

Let us set $Q = A \cdot P$, a matrix of one column and α rows, and denote its elements by $q_1, q_2 \cdots q_\alpha$. Also let $y_1, y_2 \cdots y_\beta$ be the elements of the matrix $Y = B^{-1} \cdot X$, which is of one column and β rows. Then (15) becomes:

$$E^* \cdot Y = Q, \tag{16}$$

which is equivalent to the set of α equations:

$$\begin{aligned} d_i y_i &= q_i \qquad (0 < i \leq r), \\ 0 &= q_j \qquad (r < j \leq \alpha), \end{aligned} \tag{17}$$

where r is the rank of E. If equations (17) are to be consistent, the q_j's must all be zero, and in this case the solution is.

$$\begin{aligned} y_i &= \frac{q_i}{d_i} \qquad (0 < i \leq r), \\ y_j &\text{ is arbitrary} \qquad (r < j \leq \beta). \end{aligned} \tag{18}$$

To express the condition that equations (17) be consistent and *solvable in integers*, in terms of the coefficients of (11), we proceed as follows. Form the "augmented matrix" of the system, a matrix $\overline{E}$ of α rows and $\beta + 1$ columns whose ith row has as its elements:

$$\epsilon_i^1, \epsilon_i^2, \cdots, \epsilon_i^\beta, -p_i.$$

The matrix $\overline{S}$ formed by multiplying $\overline{E}$ on the left by A will have as the elements of its ith row $(0 < i \leq \alpha)$:

$$s_i^1, s_i^2, \cdots, s_i^\beta, -q_i,$$

where the s_i^j's are the elements of the matrix:

$$S = \| s_i^j \| = A \cdot E.$$

Since multiplying the matrix S by B reduces it to the normal form, it may be reduced to the form E^* by operations on columns only; which shows that $\overline{S}$ may be reduced by operations on columns only to the form:

$$\overline{E}^* = \left\| \begin{array}{cccccccc} d_1 & 0 & \cdots & 0 & 0 & \cdots & 0 & -q_1 \\ 0 & d_2 & \cdots & 0 & 0 & \cdots & 0 & -q_2 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & d_r & 0 & \cdots & 0 & -q_r \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & -q_\alpha \end{array} \right\| . \tag{19}$$

In order that (11) be solvable at all, we found that q_i must be zero for values of i greater than r. This shows that the rank of $\bar{E}^*$ is r. If in addition we require the solutions to be integers, q_i must be divisible by d_i for $i \leq r$. Hence $\bar{E}^*$ may be reduced to normal form by adding the ith column to the last q_i/d_i times. Hence its invariant factors must be the same as the d_i's, i.e., those of E. Conversely, if this condition is satisfied, each q_i will be divisible by the corresponding d_i, and the solutions of (11) will be integers.

Since $\bar{E}^*$ was obtained from $\bar{E}$ by elementary transformations, it has the same rank and invariant factors as $\bar{E}$. Hence we have proved the two theorems:

A necessary and sufficient condition that the equations (11) *have a set of integral solutions is that the augmented matrix $\bar{E}$ have the same rank and invariant factors as the matrix of the coefficients E.*

13. Since the solutions of (17) are given by (18), and since $X = B \cdot Y$, the solutions of (11) are:

$$(20) \qquad x_i = \sum_{j=1}^{j=\beta} b_i^j y_j = \sum_{j=1}^{j=r} b_i^j \frac{q_j}{d_j} + \sum_{j=r+1}^{j=\beta} b_i^j y_j \qquad (0 < i \leq \beta),$$

in which $y_{r+1}, y_{r+2}, \cdots, y_\beta$ are arbitrary integers.

If the equations were homogeneous, the p_i's would all be zero, and hence the q_i's would also be zero. Hence the solutions would be of the form:

$$(21) \qquad x_i = \sum_{j=r+1}^{j=\beta} b_i^j y_j \qquad (0 < i \leq \beta),$$

in which $y_{r+1}, y_{r+2}, \cdots, y_\beta$ are arbitrary integers.

Consequently, for such equations we have the theorem:

A set of linear homogeneous equations whose coefficients are integers has a set of $\beta - r$ linearly independent solutions each of which is a set of relatively prime integers, if β is the number of unknowns and r the rank of the matrix of the coefficients. All other solutions in integers are linearly dependent on these $\beta - r$ linearly independent solutions, the coefficients of the linear relations being integers.

This result was to be expected, since if a set of linear homogeneous equations are solvable in rational numbers, they are solvable in integers.

By comparing (20) and (21) we obtain the further result:

If one set of integers satisfying equations (11) *be given, the other solutions are obtained by adding to it the solutions of the homogeneous equations which result when the right members of* (11) *are replaced by zeros.*

The theorems of this paragraph were first given in complete form by H. J. S. Smith,* although he was anticipated to some extent by Heger.†

---

* Smith, H. J. S., On Systems of Linear Indeterminate Equations and Congruences, Philos. Transactions, Vol. 151, pp. 293 f. Collected Works, XII, pp. 367 ff.

† Heger, Ignaz, Mem. Vienna Academy, Vol. XIV, second part, p. 111.

Skew-Symmetric Matrices.

14. A skew-symmetric matrix is one in which

$$\epsilon_i^j = -\epsilon_j^i. \tag{22}$$

Let us define as the *conjugate* of a square matrix the matrix obtained from it by interchanging rows and columns. Evidently if a skew-symmetric matrix be pre-multiplied by any square matrix, and post-multiplied by the conjugate of this matrix, it will remain skew-symmetric.

If A is the matrix defined in § 9 such that

$$A \cdot E \cdot B = E^*, \tag{6}$$

the matrix $A \cdot E$, by § 11, has d_1 as the H. C. F. of the elements of the first row. Since multiplication on the right corresponds to operations on columns only and leaves the H. C. F. of the elements of the first row unchanged, d_1 is also the H. C. F. of the elements of the first row of $A \cdot E \cdot A'$ where A' denotes the conjugate of A. Since $A \cdot E \cdot A'$ is skew-symmetric, d_1 will also be the H. C. F. of the elements of its first column.

We reduce $A \cdot E \cdot A'$ further as follows: If the second element of the first row does not divide all the remaining elements in that row, let ϵ_1^j be one which it does not divide. Subtract the second column from (or add it to) the jth a number of times equal to the greatest integer in the quotient $\epsilon_1^j/\epsilon_1^2$, thus replacing ϵ_1^j by an element numerically less than ϵ_1^2. Upon subtracting the jth column from (or adding it to) the second, we obtain an element in the first place of the second column smaller than the one there before. All these operations leave the first column unchanged, and since the matrix was skew-symmetric, a similar set of operations on the rows reduces the matrix to a skew-symmetric matrix with the first element in the second column numerically smaller than before. By repeating these operations a sufficient number of times—at most $|\epsilon_1^2|$ times—this first element will be the highest common factor of the elements in the first row, and consequently of the elements of the matrix. When this condition is reached, we combine the second column with the other columns such a number of times that all the elements in the first row after the second will be zero, and perform similar operations on the rows. Then we combine the first column with the other columns such a number of times that all the elements in the second row after the first will be zero, and perform similar operations on the rows. This will reduce our matrix to the form:

$$E_1 = \left\| \begin{array}{cccccc} 0 & d_1 & 0 & 0 & \cdots & 0 \\ -d_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \epsilon_3^4 & \cdots & \epsilon_3^\alpha \\ 0 & 0 & \epsilon_4^3 & 0 & \cdots & \epsilon_4^\alpha \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \epsilon_\alpha^3 & \epsilon_\alpha^4 & \cdots & 0 \end{array} \right\| . \tag{23}$$

The matrix obtained from E_1 by deleting its first two rows and columns is skew-symmetric, and by applying the above process to it, in a way wholly analogous to the way by which we extended our initial process of reduction in § 9, we may, by a finite number of operations, reduce our matrix to one, $\bar{E} = \| \epsilon_i^j \|$, in which

$$\epsilon_{2i-1}^{2i} = d_i; \qquad \epsilon_{2i}^{2i-1} = -d_i \qquad (0 < i \leqq p)$$

and the remaining elements are zero. That is, our matrix consists of a series of skew blocks of two non-zero elements each along the main diagonal, surrounded by zero elements.

Since in the above process we have always performed identical operations on rows and columns, we may write:

$$E = U \cdot \bar{E} \cdot U' \tag{24}$$

where U and U' are conjugate matrices whose determinants are $+1$.

Since interchanging the first and second, third and fourth, $\cdots$ $(2n - 1)$th and $2n$th rows, and changing the signs of the even rows would reduce this matrix to the usual normal form E^*, the d_i's appearing in $\bar{E}$ must be identical with those of E^*, i.e., the invariant factors of E. Hence, we have the result:

The invariant factors of a skew-symmetric matrix are equal in pairs, and the rank of such a matrix is an even number. A skew-symmetric matrix may be reduced to the " skew " normal form, $\bar{E}$, by multiplying on the left by a unimodular matrix U and on the right by its conjugate, U'.

Symmetric Matrices.

15. A *symmetric* matrix is one in which:

$$\epsilon_i^j = \epsilon_j^i. \tag{25}$$

Since a symmetric matrix retains its symmetry when we perform any operations on its rows, provided we perform the same operations on its columns, the question naturally arises whether a process similar to that of the preceding paragraph exists which will enable us to reduce such matrices to their normal form by means of a matrix and its conjugate. This question must be answered in the negative,[*] as is proved by the following example. The matrix

$$E = \left\| \begin{matrix} 2 & 1 \\ 1 & 2 \end{matrix} \right\|, \tag{26}$$

---

[*] On p. 189 of Scott and Mathews' Determinants the erroneous statement is made that symmetric matrices with integral elements can always be reduced to normal form by identical operations on rows and columns.

can not be reduced to its normal form,

$$E^* = \begin{Vmatrix} 1 & 0 \\ 0 & 3 \end{Vmatrix}, \tag{27}$$

by a matrix

$$U = \begin{Vmatrix} a & b \\ c & d \end{Vmatrix}, \tag{28}$$

and its conjugate U', where a, b, c and d are integers, since one of the conditions the elements of U would have to satisfy is:

$$2a^2 + 2ab + 2b^2 = 1. \tag{29}$$

The reduction and classification of symmetric matrices by identical operations on rows and columns is thus a problem of a different order from those which have been considered in this paper. Equivalence under such operations involves much more than the equivalence of invariant factors. This classification is the fundamental problem of the arithmetic theory of quadratic forms.*

Matrices with Elements Reduced, Modulo 2.

16. In many applications of matrices to Analysis Situs, it is found convenient to reduce the elements of the matrices modulo 2. On reducing modulo 2, the equation (6) becomes

$$\bar{A}\cdot\bar{E}\cdot\bar{B} = \bar{E}^*, \tag{30}$$

in which $\bar{E}$ can represent an arbitrary matrix of α rows and β columns whose elements are 0 and 1, $\bar{A}$ and $\bar{B}$ represent square matrices of determinant unity (mod. 2) of α and β rows respectively, and $\bar{E}$ is a matrix all of whose elements are 0 except a sequence of elements along the main diagonal which are 1. This follows from the fact that if one of the invariant factors of E^* is even, so are all the following invariant factors since they contain this one as a factor. The number of 1's in $\bar{E}^*$ is the rank of $\bar{E}^*$. It is less than or equal to the rank of E^*, and differs from it by the number of even invariant factors of E.

Symmetric Matrices, Modulo 2.

17. The theory of symmetric matrices, mod. 2, is not subject to the difficulties referred to in § 15. The reduction of such a matrix to normal form may be effected as follows: First interchange rows (performing the same interchange of columns) until the main diagonal consists of a series of 1's followed by a series of 0's. This can be effected by elementary

---

* Cf. Encyclopédie des Sciences Mathématiques, Tome I, vol. 3, p. 101.

transformations according to § 4 because the negative of any element is the same as the element itself, modulo 2. Add the first row to every row whose first element is a 1, and the first column to the corresponding columns. Repeat this for the second row and column performing a new interchange of rows and columns, if necessary, and continue until there are no elements different from zero in the main diagonal after those used.

The part of the matrix still to be normalized is now in the skew-symmetric form, since $+1 = -1$ (mod. 2) and may be normalized by the process of § 14. Thus by identical operations on rows and columns we have reduced our matrix to the form $\bar{E}^* = \| \epsilon_i^j \|$ in which:

$$\epsilon_i^i = 1 \ (0 < i \leqq p); \quad e_{p+2i-1}^{p+2i} = 1; \quad e_{p+2i}^{p+2i-1} = 1 \ (0 < i \leqq q),$$

and all the remaining elements of the matrix are zero. That is, the non-zero elements consist of a series of 1's in the main diagonal, followed by a series of skew blocks, each containing two 1's. If $p = 0$, this matrix can not be reduced further; but if $p \neq 0$, it may be reduced to a form containing one or two 1's in the main diagonal (according as p is odd or even) and a series of skew blocks, or to a form containing a series of 1's in the main diagonal and no skew blocks. This further reduction depends on the fact that a group of three 1's in the main diagonal of a matrix in the above form may be replaced by a single 1 in the main diagonal and a skew block of two.

The steps of the process in the case of a three-rowed square matrix are, first adding the first row and column to the second row and column respectively, then adding the third row and column to the first row and column respectively, and finally adding the second row and column to the third. The matrix becomes successively:

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix}.$$

These steps may be made in reverse order to effect the inverse transformation, and are obviously typical of the steps which can be applied to any matrix $\bar{E}^*$ for which $p > 0$.

Symmetric Matrices, Modulo p.

18. In the case of an odd prime modulus, the reduction of a symmetric matrix is even simpler than in the modulo 2 case, and the normal form is a matrix all of whose elements are 0 except a sequence down the main diagonal. For, by a set of interchanges on rows, followed by similar ones on columns, we can obtain a non-zero element in the first row

of the matrix. If this is not the leading element, by adding the column containing it to the first column, and the corresponding row to the first row, we obtain a leading element, ϵ_1^1, which is not zero. We may then add the first row to the others so as to make all the other elements in the first column 0, and operate similarly on the columns. The number of times, n, we must add the first row to a row with first element ϵ_r^1 is given by solving the congruence:

$$n\epsilon_1^1 + \epsilon_r^1 \equiv 0 \qquad (\text{mod. } p) \tag{31}$$

which has a root since p is prime. The reduction is continued as in § 9. By interchanges of rows and columns after we have reduced our matrix to a form similar to that given by (5), we can change the order of the d_i's, and have to decide on some definite order to get a single normal form. We may, for example, write the $p - 1$ non-zero elements of our system as the integers from 1 to $(p - 1)/2$ with plus and minus signs; and take as the normal order that of absolute value of the elements when written in this form.